

CORR. U> 6-052-666

[19]中华人民共和国专利局

[51]Int.Cl<sup>6</sup>

G06K 9/62

G10L 9/00



[12]发明专利申请公开说明书

[21]申请号 96112486.5

[43]公开日 1997年8月20日

[11]公开号 CN 1157444A

[22]申请日 96.10.30

[74]专利代理机构 柳沈知识产权律师事务所

[30]优先权

代理人 马 莹

[32]95.11.6 [33]EP[31]95402468.3

[71]申请人 汤姆森多媒体公司

地址 法国库伯瓦

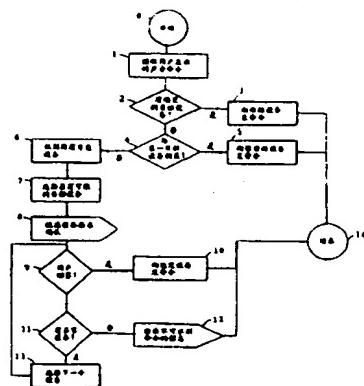
[72]发明人 埃里克·迪尔 杰勒德·科达

权利要求书 2 页 说明书 4 页 附图页数 1 页

[54]发明名称 家用设备的语音识别

[57]摘要

一种基于语音的人-机通讯系统，包括多个具有语音合成功能的可控设备，每个所述设备有其独有的声音模式，各设备通过总线连接以使中央处理机构处理所有来自用户的请求，由于用户使用其自然语言，命令会引起歧义，因此提供了一个用于处理歧义情况的算法。



Best Available Copy

(BJ)第 1456 号

## 权 利 要 求 书

- 1、带有多个具有语音合成装置的可控设备的基于语音的人-机通讯系统，其特征在于，为每个可控设备提供一种其自身独有的声音模式。
- 5 2、根据权利要求1的人-机通讯系统，其中可控设备是家用环境下的家用设备。
- 3、根据权利要求1或2的人-机通讯系统，其特征在于，该系统具有至少一个语音识别装置。
- 4、根据上述权利要求之一的人-机通讯系统，其中提供一个总线系统  
10 来相互连接所述可控设备。
- 5、根据权利要求4的人-机通讯系统，其中所述系统具有唯一的中央处理机构。
- 6、根据权利要求5的人-机通讯系统，其中所述语音识别系统在中央处理机构中。
- 15 7、根据权利要求5或6的人-机通讯系统，其特征在于，由所述中央处理机构来管理所有的请求。
- 8、根据权利要求7的人-机通讯系统，其特征在于，由所述中央处理机构根据以下步骤确定目标设备：
  - a)如果明确指出了目标设备，中央处理机构就向目标设备发一个命令；
  - 20 b)如果没有明确指出目标设备，中央处理机构就检查该命令是否仅与唯一的设备有关，
    - b1)如果 b)为真，则中央处理机构向此设备发一个命令；
    - b2)如果 b)为假，中央处理机构便产生可能理解此命令的所有设备的列表。
- 25 9、根据权利要求7的人-机通讯系统，其特征在于，根据可能目标设备的列表，每个可能目标设备由中央处理机构触发向用户发出询问直到得到一个肯定回答为止。
- 10、根据权利要求8或9的人-机通讯系统，其特征在于，可能目标设备的列表是根据一概率方法产生的。
- 30 11、在人-机通讯系统中选择一个目标设备的方法，所述人-机通讯系统在其中央处理机构部件中装有一个语音识别装置和具有多个带有独有的语

# 说 明 书

## 家用设备的语音识别

5 本发明涉及尤其用于家用的可控设备的语音识别，特别涉及到一个基于语音的人 - 机通讯系统和一种从多个设备中确定一个目标设备并与该设备进行通讯的方法。

目前对于可控设备，尤其象盒式磁带录象机、电视机和 CD 播放机这样的家用设备的控制在多数情况下是使用借助红外线来发送信息或命令的一个或多个遥控部件来执行的。选定设备向用户的反馈通常是由执行命令结果的光显示来提供的。其它的人 - 机界面是已知的。例如当计算机控制一定数量的设备时(不一定是家用设备)其中人与机器之间的界面通常由键盘、鼠标和屏幕来提供，借助这些界面中央处理器件通过程序来控制这些被连接的设备。

15 由于家用设备的日益增多及其日趋复杂，上述人 - 机界面对用户并不友好，一种增强用户界面的用户友好程度的方法是增加用户界面所需传感装置的数量。一种最有前途的自然输入/出的可行方案是语音界面。

声音既可通过语音识别用于命令，又可用于通过语音合成进行的反馈。目前使用语音合成的应用是为仅仅管理一个设备的环境而设计的。在这一已知的实例中，设备配备有一个语音识别和一个语音合成系统。这样的解决方案在象机器人这样的领域是已知的。相对照，家用设备通常包括原则上可控的多个不同设备。这就提出了如果使用基于语音的人机界面，如何在用户与设备之间的对话中识别不同可控设备的问题。

因此，本发明的一个目的是提供一种在用户与其家用设备之间建立“自然”对话的基于语音的人 - 机通讯系统。

此目的通过各独立权利要求涉及的内容来实现。本发明优选实施例由各从属权利要求说明。

声音反馈主要发生在两种情况下：

产生一项帮助用户的功能，例如当用户执行一项复杂工作时设备会提供30 指导。

当用户没有将注意力集中在有关设备上时，通知或警告用户。

在所知的单一设备环境中，由于只有一个设备发送消息，因此反馈是直接的。在多设备环境中，每个设备必须提供一个额外信息，即发送该消息的设备的识别信息。该设备的识别信息是由权利要求 1 表述特征给定的，其中对每个可控设备提供它的一个唯一的声音。换言之，每个设备都有它自己的声音合成器，此合成器能以一种方式合成该特定设备的声音以便有可能识别这一设备。这对于用户是非常友好的，因为在自然环境中，每个人是靠其自身的声音被识别的。因此，该设备传递的消息暗含着发出者的识别信息。

此外，这种语音识别可通过为每个设备提供一个与用户头脑中关于此设备的图像相适合的声音而得到增强。例如，在法国，电视机被认为是女性设备，而盒式磁带录像机被认为是男性设备。这种情况下，为电视机提供一个女性声音而为盒式磁带录像机提供一个男性声音将是有益的。并且还有可能为每个设备配备几种不同的声音，以使用户可选择较喜欢的声音。由于可以用这种方法给语音合成器编程，因此不会产生任何问题。

因此，本发明包含一个带有每个设备都有语音合成功能的多个可控设备的基于语音的人-机通讯系统，其中每个可控设备都有其独有的声音模式。

具体而言，人-机通讯系统中的可控设备是家庭环境中的家用设备，但本发明也可应用于其它环境。

在一个根据本发明的优选实例中，所有的设备通过总线相连。总线可通  
20 过多种途径实现，例如通过一种基于导线的电学模式、通过光纤、通过无线电波或通过红外线。为降低本发明基于语言的人-机通讯系统的复杂性，系  
统有唯一的中央处理机构来处理用户发出的所有请求。为此目的，中央处理  
25 机构配备有一个语音识别器来收集用户发出的所有请求，并向有关设备发出  
命令。这些用户命令能直接被象用于一个或几个设备的遥控器这样的标准输入  
设备激活或通过声音消息激活。那种情况下，通过中央处理机构直接接收  
分析声音消息。语音识别器目前正处于实验研究中，在市场上可买到现有的  
产品，这样条件限制了昂贵的语音识别硬件与软件的增长。

为获得一个用户友好的环境，声音界面要尽可能地接近自然语言。虽然  
30 这种解决方案是诱人的，但它有一个主要的弊端，主要是同样的命令可被不同的设备理解。例如短语“换到第五套节目”对电视机和盒式磁带录像机来说都是有意义的。

为避免命令的混乱，有一种简单的解决方案即由用户来明确指出他想命

音合成装置的目标设备，其中在中央处理机构和目标设备之间的通讯是在总线上进行的，其特征在于，所述中央处理机构根据以下步骤来确定所述目标设备：

- a)如果明确指出了目标设备，中央处理机构就向该目标设备发一个命令；
- b)如果没有明确指出目标设备，则中央处理机构检查该命令是否仅与唯一的设备有关；
  - b1)如果 b)为真，则中央处理机构向此设备发一个命令；
  - b2)如果 b)为假，则中央处理机构产生一个可能理解此命令的所有设备的列表。

12、根据权利要求 11 的选择一个目标设备的方法，其特征在于，可能目标设备的列表是根据一概率方法产生的。

令的设备。这种方法直接但不灵活。因此，本发明采用了一种较灵活自然的算法，现描述如下：

当语音识别器收到一个命令时，由中央处理机构检查是否明确指出了目标设备。这种情况下，命令指向被调用或指出的设备。否则，中央处理机构  
5 检查命令是否仅与一单独设备有关以便该命令被指向此单独设备。否则，中央处理机构列出可能理解此命令的所有设备。然后，每个被列出的设备用简明的语言询问用户此命令是否是对它的。这个过程一直持续到例如用户给出肯定答复来确定他所指的目标设备时为止。设备名单可通过统计或计算概率的方法产生。由于每个设备都有它自己的声音，那么用户能与设备保持自然  
10 的对话。

用于选择被调用设备的中央处理机构采用算法的优选实施例将结合附图通过实例作更为详尽的描述，其中

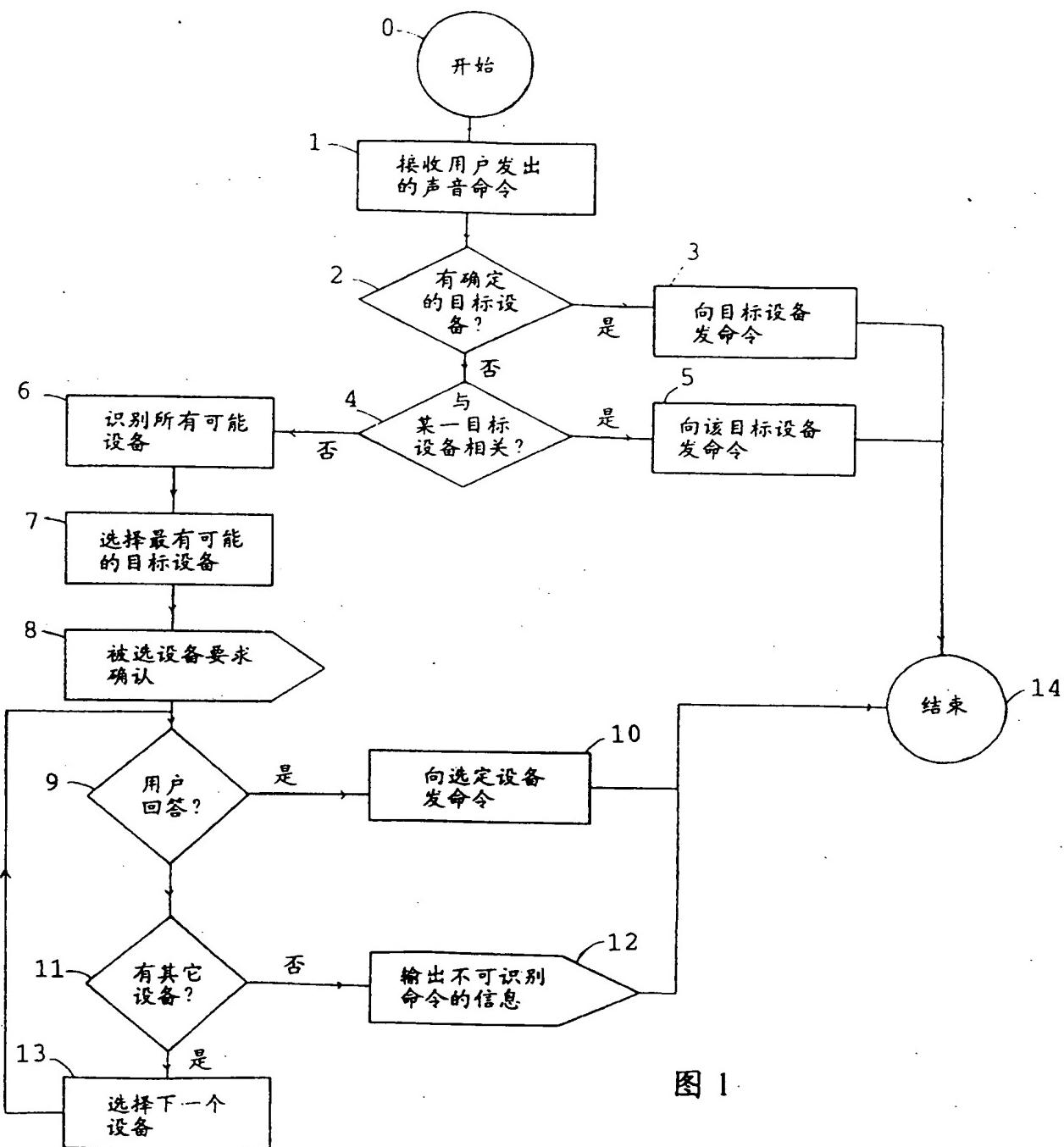
图 1 示出了所使用算法的流程图。

图 1 示出了在中央处理机构中执行的算法的流程图，中央处理机构可由适当的计算机如个人电脑或工作站形成。算法以步骤 0 开始。在步骤 1 中中央处理机构接收到一个来自用户的声音命令。步聚 2 中分析此命令以确定是否说明了目标设备。若所使用设备已正确地命名，任务就会执行下去。如果回答为“是”，步骤 3 中中央处理机构就把命令发往该目标设备，算法就结束了(步骤 14)。如果回答为“否”，中央处理机构则分析此命令并确定它是否与某一特定目标设备相关。换言之，作为相连设备的一项功能，中央处理机构检查命令是否明确。如果回答为“是”，算法接着进行将命令发往目标设备的步骤 5。之后算法结束(步骤 14)。如果回答为“否”，中央处理机构就会在步骤 6 中对所有与该给定命令相关的可能设备进行识别。在步骤 6 中，中央处理机构还产生一个可能目标设备的列表。所述列表的顺序可由如统计方法之类的方法确定或决定。在步骤 7 中，选定了最有可能的目标设备。由于中央处理机构的命令，在步骤 8 中这样选定的设备要求得到确认，即它的声音合成被激活。在步骤 9 中，中央处理机构对用户的回答进行分析。如果回答为“是”，在步骤 10 中就将命令送往选定的设备。之后算法结束(步骤 14)。如果不是肯定回答或经过一段预定时间仍无回答，中央处理机构就会在步骤  
25 30 11 中检查步骤 6 中的列表是否包含其它设备。如果列表中不包含其它设备，中央处理机构就会在步骤 12 中给用户输出一个不可识别的命令的信息，之

后算法结束。如果列表中仍有多个设备，中央处理机构就在步骤 13 中选择列表中的下一个设备，之后该设备向用户要求确认并返回步骤 9。

处理过程可以优化，既然中央处理机构已选择了所有相关的设备，它也可对其按照概率下降的顺序进行分类。例如，如果用户在电视机打开时请求 5 换节目，则该命令更有可能是对电视机而不是对盒式磁带录像机。由于这个原因，电视机首先说话。如果两台设备都是关闭的，则对每个设备的机会是均等的。

## 说 明 书 附 图



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**